Theory and Decision (2006) 60: 175–191
DOI 10.1007/s11238-005-4591-z

SEBASTIAN MAAß

A PHILOSOPHICAL FOUNDATION OF NON-ADDITIVE MEASURE AND PROBABILITY

ABSTRACT. In this paper, non-additivity of a set function is interpreted as a method to express relations between sets which are not modeled in a set theoretic way. Drawing upon a concept called "quasi-analysis" of the philosopher Rudolf Carnap, we introduce a transform for sets, functions, and set functions to formalize this idea. Any image-set under this transform can be interpreted as a class of (quasi-)components or (quasi-)properties representing the original set. We show that non-additive set functions can be represented as signed σ-additive measures defined on sets of quasi-components. We then use this interpretation to justify the use of non-additive set functions in various applications like for instance multi criteria decision making and cooperative game theory. Additionally, we show exemplarily by means of independence, conditioning, and products how concepts from classical measure and probability theory can be transfered to the non-additive theory via the transform.

KEY WORDS: conditioning, independence, Möbius transform, non-additive measure, products, quasi-analysis

1. INTRODUCTION

The starting point of our research is that we want to interpret non-additivity of a set function μ, i.e. $\mu(A \cup B) \neq \mu(A) + \mu(B)$ for disjoint sets A and B, in such a way that a) there is a relation between the sets A and B not being modeled in a set theoretic way since this would imply $A \cap B \neq \emptyset$ and that b) $\mu(A) + \mu(B) - \mu(A \cup B)$ is a measure for the strength of this relation. To formalize this idea, we draw upon a concept called "quasi-analysis" introduced by the philosopher Rudolf Carnap in "The logical construction of the world" in 1928 being a generalization of some abstraction principles of Frege, Russell and Whitehead.

Carnap's concept of quasi-analysis can be outlined as follows. Suppose there is given a set of basis elements and a system of logical connections over this set, either in the form of binary relations or in the form of a set system over the set of basis elements. In the latter case, every element of the set system represents a component or property being shared by each of its elements. Since the basis elements are supposed to be indivisible unities, these sets are called "quasi-components" or "quasi-properties" of the basis elements. This method of analyzing basis elements using their quasi-components is called quasi-analysis. If the set system is large enough then every basis element can be represented by the set of its quasi-components.

The next step is to apply the quasi-analysis to our problem. Starting with a non-empty and for the moment finite set Ω and an algebra $\mathcal{A}$ over Ω, we interpret sets $A \in \mathcal{A}$ in two different ways, either just as a set containing its elements or as a quasi-component of its elements. In the latter case A stands for a quasi-component or quasi-property being shared by all basis elements of A and not being shared by all basis-elements outside A. For every set $A \in \mathcal{A}$ of basis elements, we call a set $B \in \mathcal{A}$ a "quasi-component of A", if B is a quasi-component of some element of A, i.e. if $B \cap A \neq \emptyset$. Therefore, any set A of basis elements can be represented by the set $\widehat{A}$ of its quasi-components, $\widehat{A} := \{B \in \mathcal{A} | B \cap A \neq \emptyset\}$. After canonically extending this definition of $\widehat{A}$ to the general case with not necessarily finite Ω, we show that any set function $\mu\colon \mathcal{A} \to \mathbb{R}$ can be transformed into a (signed) measure $\widehat{\mu}$ on the σ-algebra $\widehat{\mathcal{A}} := \sigma\{\widehat{A} | A \in \mathcal{A}\}$ generated by the sets of quasi-components of sets in $\mathcal{A}$ such that $\widehat{\mu}$ preserves the values of μ, i.e. $\widehat{\mu}(\widehat{A}) = \mu(A)$ holds fox all $A \in \mathcal{A}$. Thus, μ assigns to a set $A \in \mathcal{A}$ the aggregated value of its quasi-components or, in other words, μ can be decomposed into the $\widehat{\ }$-operator mapping sets $A \in \mathcal{A}$ to their representing sets $\widehat{A}$ of quasi-components and the measure $\widehat{\mu}$ defined on sets of quasi-components. Furthermore, every μ-(Choquet-)integrable function $f\colon \Omega \to \mathbb{R}$ can be transformed into a $\widehat{\mu}$-integrable function $\widehat{f}$ on the set of quasi-components such that the value of the μ-integral is preserved, i.e. $\int \widehat{f}\, \mathrm{d}\widehat{\mu} = \int f \mathrm{d}\mu$ holds for all f. As one simple result, we obtain $\widehat{\mu}(\widehat{A} \cap \widehat{B}) = \mu(A) + \mu(B) - \mu(A \cup B)$ solving our problem introduced at the beginning as (a) $\widehat{A} \cap \widehat{B}$ represents all those quasi-components shared by A and B and (b) $\widehat{\mu}(\widehat{A} \cap \widehat{B})$ is the measure for the strength of this relation.

Although, the introduced transform of sets, functions, and set functions is mathematically almost a reformulation of the Möbius transform, its interpretation basing upon quasi-components is unique and has two important implications. First, the decomposition of a set function μ into the $\widehat{\ }$-operator defined on $\mathcal{A}$ and the measure $\widehat{\mu}$ can be used to justify the use of non-additive set functions in various theories of mathematical economics and finance as quasi-components naturally appear there even though sometimes only implicitly. This will exemplarily be shown in Section 4 by means of multi criteria decision making and cooperative game theory. Second, whenever the interpretation of the existence of quasi-components is appropriate when using non-additive probability theory then probabilistic concepts like independence and conditioning can be canonically defined via the presented transform. This will be done in Section 5. Even though our definitions of independence, conditioning, and products are different from already existing ones, they are well-founded as they are based upon a very comprehensive interpretation of non-additive set functions.

2. NOTATIONS

Throughout this paper, let Ω denote a non-empty set, $\mathcal{A}$ a σ-algebra over Ω, and μ: $\mathcal{A} \to [0, \infty[$ a set function on $\mathcal{A}$ with $\mu(\emptyset) = 0$. For every set function μ its **dual** set function $\overline{\mu}$ is defined by $\overline{\mu}(A) := \mu(\Omega) - \mu(A^c)$. A set function μ is called **monotone** if $A \subset B$ implies $\mu(A) \leqslant \mu(B)$. It is called k-**monotone** if

$$\mu\left(\bigcup_{i=1}^{k} A_i\right) + \sum_{\substack{I\subset\{1,\dots,k\}\\ I\neq\emptyset}} (-1)^{|I|}\mu\left(\bigcap_{i\in I} A_i\right) \geqslant 0$$

for $k \geqslant 2$ and $A_1, \dots, A_k \in \mathcal{A}$,

k-**alternating** if $\overline{\mu}$ is k-monotone or, equivalently, if

$$\mu\left(\bigcap_{i=1}^{k} A_i\right) + \sum_{\substack{I\subset\{1,\dots,k\}\\ I\neq\emptyset}} (-1)^{|I|}\mu\left(\bigcup_{i\in I} A_i\right) \leqslant 0$$

for $k \geqslant 2$ and $A_1, \dots, A_k \in \mathcal{A}$,

totally monotone if μ is monotone and k-monotone for every $k \geqslant 2$, and **totally alternating** if $\overline{\mu}$ is totally monotone or, equivalently, if μ is monotone and k-alternating for every $k \geqslant 2$. A **belief function** is a totally monotone set function μ being normalized, i.e. $\mu(\Omega) = 1$. A **plausibility function** is a totally alternating, normalized set function.

2-monotone set functions are also called **supermodular** or **convex** and 2-alternating set functions are also called **submodular** or **concave**. A simple example of totally monotone set functions being used in the sequel is a **unanimity game** u_A of a set $A \in \mathcal{A}$ defined by $u_A(B) := 1$ if $B \supset A$ and 0 else. For the dual unanimity game $\overline{u_A}$ holds $\overline{u_A}(B) = 1$ if $B \cap A \neq \emptyset$ and 0 else.

Finally, the (asymmetric) **Choquet integral** w.r.t. a finite monotone set function μ is defined by

$$\int f \ \mathrm{d}\mu := \int_{-\infty}^{0} \mu(f \geqslant x) - \mu(\omega) \ \mathrm{d}x + \int_{0}^{\infty} \mu(f \geqslant x) \ \mathrm{d}x.$$

3. THE TRANSFORM OF SETS, FUNCTIONS, AND SET FUNCTIONS

In the first section, we have seen that every element $w \in \Omega$ can be represented by the set of its quasi-components, $\{\mathcal{A} \in \mathcal{A} | \mathcal{A} \ni w\}$. Extending this introductory idea, we represent a set $A \subset \Omega$ by the class of all quasi-components of its elements,

$$\widehat{A} := \{B \in \mathcal{A} | B \cap A \neq \emptyset\}. \qquad (1)$$

Furthermore, we denote by $\widehat{\mathcal{A}}$ the σ-algebra generated by all sets representing classes of quasi-components,

$$\widehat{\mathcal{A}} := \sigma\{\widehat{A} \mid A \in \mathcal{A}\}. \tag{2}$$

In the subsequent proposition, we collect elementary results on the transformed sets. We omit the proof since it consists of simple calculations.

PROPOSITION 3.1. *For* $A, B \in \mathcal{A}$

(a) $\widehat{\emptyset} = \emptyset, \widehat{\Omega} = \mathcal{A} \setminus \{\emptyset\}$,
(b) $A \subset B \Leftrightarrow \widehat{A} \subset \widehat{B}$,
(c) $\widehat{A \cup B} = \widehat{A} \cup \widehat{B}$,
(d) $\widehat{A \cap B} \subset \widehat{A} \cap \widehat{B}$,
(e) $\widehat{A \setminus B} \supset \widehat{A} \setminus \widehat{B}$,
(f) $\hat{A}^c \subset \widehat{A^c}$.

Remark 3.2. The following statements on the relation between the $\widehat{\ }$-operator and monotone sequences can easily be proved and will be of interest later on.

$$(A_n)_{n \in \mathbb{N}} \nearrow A \Rightarrow \bigcup_{n \in \mathbb{N}} \widehat{A_n} = \widehat{\bigcup_{n \in \mathbb{N}} A_n}, \tag{3}$$

$$(A_n)_{n \in \mathbb{N}} \searrow A,\ A_n \neq A \forall n \in \mathbb{N} \Rightarrow \bigcap_{n \in \mathbb{N}} \widehat{A_n} \supsetneqq \widehat{\bigcap_{n \in \mathbb{N}} A_n}. \tag{4}$$

Especially, in the latter case, $A^c \in \bigcap \widehat{A}_n \setminus \widehat{\bigcap A_n}$.

We now address the problem wether every set-function $\widehat{\mu}$: $\mathcal{A} \rightarrow \mathbb{R}$ can be transformed into a measure $\mu: \widehat{\mathcal{A}} \rightarrow \mathbb{R}$ on the set of quasi-components, satisfying

$$\widehat{\mu}(\widehat{A}) = \mu(A) \quad \text{for all } A \in \mathcal{A}, \tag{5}$$

i.e. the measure of the set A should equal the measure of all of its quasi-components. Such a transformed set function would solve our task for interpreting non-additivity as a method of expressing a connection between the sets in $\mathcal{A}$ which is not modeled in a set theoretic way. For two disjoint sets A and B, we interpret $\mu(A)+\mu(B)-\mu(A\cup B)$ as a measure of their connection. This interpretation perfectly fits with that of $\widehat{A}\cap\widehat{B}$ being the set of all common quasi-components of A and B and of $\widehat{\mu}(\widehat{A}\cap\widehat{B})$ being a measure of their weight because

$$\begin{aligned}\widehat{\mu}(\widehat{A}\cap\widehat{B}) &= \widehat{\mu}(\widehat{A})+\widehat{\mu}(\widehat{B})-\widehat{\mu}(\widehat{A}\cup\widehat{B})\\ &= \widehat{\mu}(\widehat{A})+\widehat{\mu}(\widehat{B})-\widehat{\mu}(\widehat{A\cup B})\\ &= \mu(A)+\mu(B)-\mu(A\cup B).\end{aligned}$$

Furthermore, equation (5) can be interpreted in such a way that μ assigns to a set $A\in\mathcal{A}$ the aggregated value of its quasi-components.

According to Remark 3.2, the transform $\widehat{\mu}$ of μ won't be σ-additive in general since whenever $\lim_{n\to\infty}\mu(A_n)<\mu(A)$ for a monotone increasing sequence $(A_n)_{n\in\mathbb{N}}$ we also would have that $\lim_{n\to\infty}\widehat{\mu}(\widehat{A_n})<\widehat{\mu}(\widehat{A})$ by equations (3) and (5), i.e. $\widehat{\mu}$ wouldn't be continuous from below. In order to make $\widehat{\mu}$ σ-additive, we have to enlarge the set of quasi-components to that effect that equation (3) changes to

$$(A_n)_{n\in\mathbb{N}}\nearrow A,\ A_n\neq A\,\forall n\in\mathbb{N}\Rightarrow\bigcup_{n\in\mathbb{N}}\widehat{A_n}\subsetneqq\widehat{\bigcup_{n\in\mathbb{N}}A_n}, \tag{6}$$

while maintaining equation (4). A natural approach is very similar to the completion of the rational numbers via Cauchy sequences: The set of quasi-components (more precisely, the generator of $\widehat{\mathcal{A}}$) is then the set of equivalence classes of monotonously increasing sequences in $\mathcal{A}$, under the equivalence relation $(A_n)\sim(B_n)$ if there exists a monotonously increasing interleave sequence, i.e. if

$$\begin{aligned}(A_n)\sim(B_n)\ :\Leftrightarrow\ \exists(C_n)\text{ increasing with }&\{C_n|n\in\mathbb{N}\}\\ &=\{A_n|n\in\mathbb{N}\}\cup\{B_n|n\in\mathbb{N}\}.\end{aligned}$$

We state without a proof that every such equivalence class $[(A_n)]$ can uniquely be represented by a plausibility function:

$$\nu_{[(A_n)]}(A):=\begin{cases}1, & \text{if }\exists(B_n)\in[(A_n)]:A=B_n\quad\text{for some }n\in\mathbb{N},\\ 0 & \text{else.}\end{cases}$$

we therefore redefine

$$\widehat{A} := \{\nu \text{ plausibility function } | \nu(A) = 1\} \quad \text{for all } A \in \mathcal{A}, \tag{7}$$

$$\widehat{\mathcal{A}} := \sigma\{\widehat{A} | A \in \mathcal{A}\}. \tag{8}$$

We call the elements of any set $\widehat{A}$ the set of quasi-components of A. It is easy to prove that Proposition 3.1 remains valid and that inequalities (4) and (6) hold under the redefined terms. Moreover, it can be shown that in the discrete case any plausibility function is a dual unanimity game which again can be represented by a set in $\mathcal{A}$ such that the definitions (1) and (7) as well as (2) and (8) coincide.

We now show how μ-integrable functions f on Ω have to be transformed to $\widehat{\mu}$-integrable functions on the set $\widehat{\Omega}$ of quasi-components implicitly understanding that there exists a transform between non-additive set function and signed measures satisfying equation (5). The following equations hold for any definition of $\widehat{f}$ and finite μ,

$$\begin{aligned}
\int_\Omega f \, \mathrm{d}\mu &= \int_0^\infty \mu(f \geqslant x) \, \mathrm{d}x + \int_{-\infty}^0 \mu(f \geqslant x) - \mu(\Omega) \, \mathrm{d}x \\
&= \int_0^\infty \widehat{\mu}(\widehat{\{f \geqslant x\}}) \, \mathrm{d}x + \int_{-\infty}^0 \widehat{\mu}(\widehat{\{f \geqslant x\}}) - \widehat{\mu}(\widehat{\Omega}) \, \mathrm{d}x, \\
\int_{\widehat{\Omega}} \widehat{f} \, \mathrm{d}\widehat{\mu} &= \int_0^\infty \widehat{\mu}(\{\widehat{f} \geqslant x\}) \, \mathrm{d}x + \int_{-\infty}^0 \widehat{\mu}(\widehat{f} \geqslant x) - \widehat{\mu}(\widehat{\Omega}) \, \mathrm{d}x.
\end{aligned}$$

In order to obtain $\int_\Omega f \ \mathrm{d}\mu = \int_{\widehat{\Omega}} \widehat{f} \ \mathrm{d}\widehat{\mu}$, we must guarantee $\widehat{\{f \geqslant x\}} = \{\widehat{f} \geqslant x\}$. Since

$$\nu \in \widehat{\{f \geqslant x\}} \quad \Longleftrightarrow \quad \nu(f \geqslant x) = 1 \quad \Longleftrightarrow \quad \int_\Omega f \, \mathrm{d}\nu \geqslant x,$$

we therefore define $\widehat{f}\colon \widehat{\Omega} \to \mathbb{R}$ by

$$\widehat{f}(\nu) := \int_\Omega f \, \mathrm{d}\nu \quad \text{for all } \nu \in \widehat{\Omega} \tag{9}$$

for any μ-integrable $f\colon \Omega \to \mathbb{R}$ and obtain $\widehat{\{f \geqslant x\}} = \{\widehat{f} \geqslant x\}$, and hence the desired result.

Having now fixed the base notations in (7)–(9), we proceed with stating the main theorem.

THEOREM 3.3.

(a) *For any set function* μ *on an algebra* $\mathcal{A}$ *there exists a unique* (*signed*) *measure* $\widehat{\mu}$ *on* $\widehat{\mathcal{A}}$ *satisfying* $\widehat{\mu}(\widehat{A}) = \mu(A)$ *for all* $A \in \mathcal{A}$.
(b) *The measure* $\widehat{\mu}$ *is non-negative if and only if* μ *is totally alternating.*
(c) *For any* μ*-integrable function* f *holds* $\int f \, \mathrm{d}\mu = \int \widehat{f} \, \mathrm{d}\widehat{\mu}$.

Theorem 3.3 states in terms of quasi-analysis that measuring a set A with μ equals measuring the set $\widehat{A}$ of all quasi-components of A with $\widehat{\mu}$. We will see in the next section that this interpretation is not just an abstract academic gimmick as it can serve as a justification for the introduction of non-additivity in various fields of mathematical economics and finance.

The proof of Theorem 3.3 bases on some results of the Möbius transform which, now recapitulate in terms introduced by Denneberg (1997, Theorems 5.2 and 6.2, Proposition 5.3 and Example 4.1). Further noteworthy papers on this topic are by Chateauneuf and Jaffray (1998), Gilboa and Schmeidler (1995), and Marinacci (1996).

THEOREM 3.4. *For any* $A \in \mathcal{A}$ *and integrable* f *let* $\widetilde{A}, \widetilde{\mathcal{A}}$, *and* $\widetilde{f}$ *be defined by*

$$\widetilde{A} := \{\nu \text{ belief function} \,|\, \nu(A) = 1\}, \tag{10}$$

$$\widetilde{\mathcal{A}} := \sigma\{\widetilde{A} | A \in \mathcal{A}\}, \tag{11}$$

$$\widetilde{f}(\nu) := \int f \, \mathrm{d}\nu \text{ for all } \nu \in \widetilde{\Omega}, \tag{12}$$

(a) *For any set function* μ *on an algebra* $\mathcal{A} \subset 2^{\Omega}$ *there exists a unique* (*signed*) *measure* $\widetilde{\mu}$ *on* $\widetilde{\mathcal{A}}$ *satisfying* $\widetilde{\mu}(\widetilde{A}) = \mu(A)$ *for all* $A \in \mathcal{A}$. *The transformed* $\widetilde{\mu}$ *of* μ *is called the* ***Möbius transformed*** *of* μ.

(b) *The measure* $\widetilde{\mu}$ *is non-negative if and only if* μ *is totally monotone.*
(c) *For any* μ*-integrable function* f *holds* $\int \widetilde{f}\ \mathrm{d}\mu = \int \widetilde{f}\ \mathrm{d}\widetilde{\mu}$.

The subsequent proof of Theorem 3.3 shows that it is mathematically almost a reformulation of Theorem 3.4. However, the Möbius transform as well as other transforms between non-additive and (σ-)additive set functions cannot be interpreted of being defined on quasi-components or quasi-properties. Precisely this interpretation will be used in Section 4 to justify the use of non-additive set functions in different applications and in Section 5 to introduce concepts from classical measure and probability theory in the non-additive theory.

Proof of Theorem 3.3. Let $T: \widetilde{\Omega} \to \widehat{\Omega}$ be defined by $T(\nu) := \overline{\nu}$. Simple calculations yield $T^{-1}(\widehat{A}) = \widetilde{\Omega} \backslash \widetilde{A^c}$, i.e. T is measurable and transforms the generating system of $\widehat{\mathcal{A}}$ into that of $\widehat{\mathcal{A}}$ and thus, the domain of the image measure $\widetilde{\mu}^T$ of $\widetilde{\mu}$ under T is $\widehat{\mathcal{A}}$. Moreover,

$$\widetilde{u}^T(\widehat{A}) = \widetilde{\mu}(\widetilde{\Omega} \backslash \widetilde{A^c}) = \widetilde{\mu}(\widetilde{\Omega}) - \widetilde{\mu}(\widetilde{A^c}) = \mu(\Omega) - \mu(A^c) = \overline{\mu}(A).$$

(a) Existence and uniqueness follows from Theorem 3.4(a) for $\widehat{\mu} := \widetilde{\widetilde{\mu}}^T$.
(b) Using Theorem 3.4(b), we obtain

$$\begin{aligned} \mu \text{ is totally alternating} &\iff \overline{\mu} \text{ is totally monotone} \\ &\iff \widetilde{\widetilde{\mu}} \text{ is non-negative} \\ &\iff \widehat{\mu} \text{ is non-negative} \end{aligned}$$

(c) By Theorem 3.4(c),

$$\begin{aligned} \int f\,\mathrm{d}\mu &= -\int -f\,\mathrm{d}\overline{\mu} = -\int \widetilde{-f}\,\mathrm{d}\,\widetilde{\widetilde{\mu}} = \int \widehat{f} \circ T\,\mathrm{d}\,\widetilde{\widetilde{\mu}} = \int \widehat{f}\,\mathrm{d}\widetilde{\widetilde{\mu}}^T \\ &= \int \widehat{f}\,\mathrm{d}\widehat{\mu}. \end{aligned} \qquad \square$$

In the discrete case, $\widehat{\mu}(\widehat{A}) = \mu(A)$ can be rewritten to

$$\mu(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

for all $A \in \mathcal{A}$ with $m: \mathcal{A}\backslash\{\emptyset\} \to \mathbb{R}, m(B) := \widehat{\mu}(\{B\})$ for all $B \in \mathcal{A}$. This result is well-known in the Dempster–Shafer theory of evidence as the representation of a plausibility functions by a "basic probability assignment" m (cf. Shafer, 1976).

Murofushi and Sugeno (1989, 1991) arrive in at representation theorems similar to Theorem 3.3 by proposing that non-additive set functions express with their non-additivity interaction among subsets which is very similar to our interpretation. But in contrast to the approach presented in this paper, they formalize this idea with an operator on $\mathcal{A}$ not preserving unions of sets (cf. Proposition 3.1(c)), interpret this as the existence of a feature that the union but neither of the sets involved has, and call it a "cooperative action" of these sets (cf. Murofushi and Sugeno, 1989 p. 206). Hence, the transformed sets can only be interpreted as quasi-properties of the original sets in situations where it is meaningful to talk about cooperative action, e.g. sometimes in cooperative game theory but never in multi-criteria decision-making. As cooperative interaction is incompatible with the interpretation of quasi-components, their transformed sets cannot represent components of the original sets. The technical difference between the two approaches is that Murofushi's and Sugeno's transform yield a σ-additive measure being non-negative while ours is signed.

4. EXAMPLES AND APPLICATIONS

In this section, we provide some examples and applications showing that the presented transform is nearly all-purpose in those theories making use of non-additive set functions and their corresponding Choquet integrals. We exemplarily show by means of multi-criteria decision-making and cooperative game theory where quasi-components naturally appear and therefore provide a justification for the use of non-additive set functions in the respective theories.

4.1. *Multi criteria decision making*

Usually, a multi criteria decision problem consists of at least two different alternatives and at least two different criteria or attributes of the alternatives. A standard method to solve such a problem is the use of aggregation: each alternative is represented by function on the set of criteria such that its values represent how "good" this criteria is distinct in this alternative. One has to introduce a weight function (measure) on the set of criteria and the rank problem can then be solved just by calculating the integral of each alternative representing function and rank them by their values. To proceed in this way, one has to impose that the criteria are unrelated since otherwise the common parts of different criteria are multiple weighted. Naturally, this presupposition is not fulfilled in general like e.g. in the following evaluation problem (cf. Grabisch, 1995 p. 295, 1996, p. 451).

Ranking students (alternatives) on the basis of their marks in different subjects (criteria) poses the problem that the different subjects do not measure mutually disjoint skills. For instance, some mathematical skills are not only tested in mathematics but also in physics. How can the students be ranked in a reasonable way incorporating these relations? The simplest approach, i.e. calculating weighted sums (i.e. expected values) of the marks is not an appropriate method as mentioned above.

Instead of an expected value as an evaluation functional, Grabisch proposes in Grabisch (1995, 1996) the use of the Choquet integral. The measurable space then consists of the set of all subjects together with an appropriate σ-algebra (usually the power set). A non-additive set function μ is able to represent correlations between subjects. For example, to model that there are skills carrying weight for the marks in mathematics and physics one has to guarantee that $\mu(\{\text{mathematics, physics}\}) < \mu(\{\text{mathematics}\}) + \mu(\{\text{physics}\})$. Roughly speaking, the more correlations between the subjects exists the more has the set function μ deviate from (σ-)additivity. In this method of resolution, the missing richness of the set of subjects to model correlations in a set theoretic way is compensated by the use of non-additive set functions.

Grabisch uses non-additive set functions to cope with skills being tested in different subjects and that cannot be modeled in a set theoretic way once he decided to model the subjects as formally indivisible unities which is not unnatural as the marks are only given to the subjects and are not decomposed to singular skills. Hence, those skills being common to different subjects have to be modeled as quasi-components of the subjects. By claiming that these common skills have to be incorporated in the valuation, Grabisch implicitly argues that μ is intrinsically defined on sets of quasi-components of the subjects but the value is assigned to the subjects. Theorem 3.3 now states that Grabisch's reasoning basing on what we called quasi-components cannot lead to anything else than non-additive set function as a solution to the multi criteria decision problem. Therefore, Theorem 3.3 justifies Grabisch's solution for the stated problem.

We now show how one explicitly arrives at Grabisch's solution by identifying the quasi-components. Let $\Omega := \{M, P, L\}$ be the set of subjects (mathematics, physics, literature). The transformed subjects together with their quasi-components are mapped in the following diagram.

The subject mathematics then consists of four quasi-components: the component $\{M\}$ of skills exclusively being necessary for mathematics, the component $\{M, P\}$ of skills being necessary for mathematics and physics but not necessary for literature and so on. To model that there is a set of skills having some influence on both marks, mathematics as well as physics, a positive value must be assigned to the set $\widehat{M} \cap \widehat{P} = \{\{M, P\}, \{M, P, L\}\}$. After such an assignment of weights to all sets of skills one directly arrives at a non-additive set function on 2^{Ω} by Theorem 3.3 as we claimed in the beginning.

4.2. *Cooperative game theory*

We now show how to apply our new transform in cooperative game theory. Let us briefly recall the setting in cooperative game theory. Denote by

S	a non-empty set of players,
$\mathcal{A} \subset 2^S$	an algebra over S interpreted as the set of all possible coalitions of players in S,

$v: \mathcal{A} \to \mathbb{R}_+$ the characteristic function of the game satisfying $v(\emptyset) = 0$ interpreted as the maximum utility/payoff the coalition A can get without correlating strategies with the other $S \backslash A$ players. This interpretation justifies the often imposed condition of v being superadditive.

In most cases, the characteristic function v is non-additive which is justified by the existence of cooperation possibilities between the coalitions which again can be interpreted as an overlap of the worth generating abilities (the quasi-components) of the coalitions. For instance, $v(C_l) + v(C_2) - v(C_1 \cup C_2) > 0$ or, equivalently, $\widehat{v}(\widehat{C_1} \cap \widehat{C_2}) > 0$, means that the coalitions C_1 and C_2 share some of their worth gaining skills. Conversely, $v(C_1) + v(C_2) - v(C_1 \cup C_2) < 0$ resp. $\widehat{v}(\widehat{C_1} \cap \widehat{C_2}) < 0$ could be interpreted to that effect that the coalitions C_1 and C_2 share some properties baring them from gaining worth or that there is a cooperative action of C_1 and C_2 (cf. Murofushi and Sugeno, 1989, p. 206). As in the preceding subsection, v is intrinsically not measuring elements of its domain but rather the set of their representing quasi-components. Again, Theorem 3.3 states that the interpretation which bases on quasi-components and the formalization with non-additive set functions match and therefore justifies the latter.

5. MATHEMATICAL CONSEQUENCES

Whenever the interpretation of the existence of quasi-components is appropriate when using non-additive probability theory then probabilistic concepts like independence and conditioning can be canonically defined via the presented transform. For example, two sets will be called independent if the two classes of quasi-components representing these sets are independent w.r.t. the transformed measure. Furthermore, also the notion of a product of non-additive set functions is introduced via the transform which is mathematically almost a reformulation of the Möbius product but it is based on a well-founded interpretation.

DEFINITION 5.1. Let $(\Omega, \mathcal{A})$ be a measurable space and $\mu: \mathcal{A} \to [0,1]$ be a set function interpreted as a not necessarily σ-additive probability measure. We call a family $(\mathcal{A}_i)_{i\in I}$, I non-empty index-set, of sub-σ-algebras of $\mathcal{A}$ **independent** if the family $(\widehat{\mathcal{A}}_i)_{i\in I}$ is independent w.r.t. $\widehat{\mu}$, i.e. if for any finite non-empty subset J of I and any sets $A_j \in \mathcal{A}_j$, $j \in J$,

$$\widehat{\mu}\left(\bigcap_{j\in J} \widehat{A}_i\right) = \prod_{j\in J} \widehat{\mu}(\widehat{A}_j).$$

Since, by the inclusion exclusion principle, $\widehat{\mu}(\bigcap_{j\in J} \widehat{A}_i) = \sum_{\emptyset\neq K\subset J} (-1)^{|K|+1} \widehat{\mu}(\widehat{\bigcup_{k\in K} A_k})$, we obtain that independence of the family $(\mathcal{A}_i)_{i\in I}$ is equivalent to

$$\sum_{\emptyset\neq K\subset J} (-1)^{|K|+1} \mu\left(\bigcup_{k\in K} A_k\right) = \prod_{j\in J} \mu(A_j) \tag{13}$$

for any non-empty finite subset J of I.
A family $(X_i)_{i\in I}$ of random variables is called **independent** if the family $(\sigma(X_i))_{i\in I}$ is independent, i.e. if

$$\sum_{\emptyset\in K\subset J} (-1)^{|K|+1} \mu\left(\bigcup_{k\in K} \{X_k \in B_k\}\right) = \prod_{j\in J} \mu(X_k \in B_k) \tag{14}$$

and any sets $B_k \in \mathcal{B}(\mathbb{R})$.

DEFINITION 5.2. Given a non-additive probability μ and two events A and B with $\mu(B) \neq 0$ the **conditional non-additive probability** of A given B is defined as the conditional probability of $\widehat{A}$ given $\widehat{B}$,

$$\begin{aligned} \nu(A|B) &:= \widehat{\mu}(\widehat{A}|\widehat{B}) \\ &= \frac{\widehat{\mu}(\widehat{A}\cap\widehat{B})}{\widehat{\mu}(\widehat{B})} \\ &= \frac{\mu(A) + \mu(B) - \mu(A\cup B)}{\mu(B)}. \end{aligned} \tag{15}$$

There are other, different approaches to introduce probabilistic concepts in non-additive probability theory (cf. Denneberg, 1994 for an overview), some of them being only an attempt to generalize the classical definitions directly where it is possible. Our approach takes into account that in some situations non-additivity shows up as a result of modeling some relations not being modeled in a set theoretic way such that the natural definitions have to be done via the introduced transform.

For the definition of products of non-additive set function we are confronted with the problem that we can not define $\mu_1 \otimes \mu_2$ via $\widehat{\mu}_1 \otimes \widehat{\mu}_2$ since the domains are not compatible. Nevertheless, it is possible to define a product in the discrete case via the transform.

DEFINITION 5.3. Let μ_1 and μ_2 be two non-additive set functions on a discrete measurable space $\Omega, \mathcal{A}$. Then their product $\mu_1 \otimes \mu_2$ is defined by

$$\mu_1 \otimes \mu_2(A) := \sum_{\emptyset \neq K_i \in \mathcal{A}_i} \widehat{\mu_1}(\widehat{K_1}) \cdot \widehat{\mu_2}(\widehat{K_2}) \cdot \overline{u_{K_1 \times K_2}}(A). \tag{16}$$

For the general case, there exist some results for the Möbius transform (cf. e.g. Brüning, 2003) which should also be able to be transfered to our setting.

6. CONCLUSIONS AND OUTLOOK

We have seen that non-additivity of set functions can often be explained by the interpretation that not all relations between sets have been modeled in a set theoretic way. Drawing upon quasi-analysis, we have introduced a transform for sets, functions, and set functions to cope with these relations. Since any image-set $\widehat{A}$ is interpreted to represent the class of the (quasi-) components of A, any result on the transformed space can directly be interpreted in terms of the original space. The use of non-additive set functions therefore does not necessarily mean a generalization of (σ-)additive measures as it can just mean that the domain was chosen too small to express all possible relations that have to be modeled. We have exemplarily shown by means of multi criteria decision making and cooperative game theory that quasi-components have a distinct meaning in different theories of mathematical economics and finance and that they are essential to justify the use of non-additive set functions in the respective theories. Moreover, the transform and its interpretation has enabled us to define independence, conditioning, and products for non-additive set functions in a canonical way. The first results arising from this new approach give reason to hope that the presented transform can establish a basis to build up a non-additive probability theory in a very natural way.

ACKNOWLEDGEMENT

Financial support by Credit Suisse and the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK) is gratefully acknowledged. Only the author is responsible for the contents of this paper.

REFERENCES

Brüning, M. (2003), Products of Monotone Measures, Möbius Transform and k-monotonicity. Dissertation thesis, Universität Bremen.

Carnap, R. (1928), *Der logische Aufbau der Welt*, Meiner, Hamburg.

Chateauneuf, A. and Jaffray, J.-Y. (1989), Some characterization of lower probabilities and other monotone capacities through the use of Möbius inversion, *Mathematical Social Sciences* 17, 263–283.

Choquet, G. (1953–1954), Theory of Capacities, *Annales de l'Institut Fourier* 5, 131–295.

Denneberg, D. (1994, 1997 2nd. ed.), *Non-additive Measure and Integral* Kluwer, Dordrecht.

Denneberg, D. (1994), Conditioning (updating) non-additive measures, *Annals of Operations Research* 52, 21–42.

Denneberg, D. (1997), Representation of the Choquet integral with the σ-additive Möbius transform, *Fuzzy Sets and Systems* 92, 139–156.

Denneberg, D. (2000), Totally monotone core an products of monotone measures, *International Journal of Approximate Reasoning* 24, 273–281.

Denneberg, D. (2002), Conditional expectation for monotone measures, the discrete case, *Journal of Mathematical Economics* 37, 105–121.

Gilboa, I. and Schmeidler, D. (1989), Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics* 18, 141–153.

Gilboa, I. and Schmeidler, D. (1995), Canonical representation of set functions, *Mathematics of Operations Research* 20, 197–212.

Grabisch M. (1995), Fuzzy integral in multicriteria decision making, *Fuzzy Sets and Systems* 69, 279–298.

Grabisch, M. (1996), The application of fuzzy integrals in multicriteria decision making, *European Journal of Operational Research* 89, 445–456.

Hendon, E., Jacobsen, H.J., Sloth, B. and Tranæs, T. (1996), The product of capacities and belief functions, *Mathematical Social Sciences* 32, 95–108.

Koshevoy, G.A. (1998), Distributive lattices and products of capacities, *Journal of Mathematical Analysis and Applications* 219, 427–441.

Marinacci, M. (1996), Decomposition and representation of coalitional games, *Mathematics of Operations Research* 21, 1000–1015.

Murofushi, T. and Sugeno, M. (1989), An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure, *Fuzzy Sets and Systems* 29, 201–227.

Murofushi, T. and Sugeno, M. (1991), A theory for fuzzy measures: representations, the Choquet integral, and null sets, *Journal of Mathematical Analysis and Applications* 159, 535–549.

Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ.

Walley, P. and Fine, T.L. (1982), Towards a frequentist theory of upper and lower probability, *Annals of Statistics* 10, 741–761.

Address for correspondence: Sebastian Maaß, Departement Mathematik, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland E-mail: sebastian.maass@math.ethz.ch